AIP Journal of Mathematical Physics

On the geometry of nonholonomic Lagrangian systems

Manuel de León and David M. de Diego

Citation: Journal of Mathematical Physics **37**, 3389 (1996); doi: 10.1063/1.531571
View online: http://dx.doi.org/10.1063/1.531571
View Table of Contents: http://scitation.aip.org/content/aip/journal/jmp/37/7?ver=pdfcov
Published by the AIP Publishing

On the geometry of non-holonomic Lagrangian systems

Manuel de León[a)]
Instituto de Matemáticas y Física Fundamental, Consejo Superior de Investigaciones Cientificas, Serrano 123, 28006 Madrid, Spain

David M. de Diego[b)]
Departamento de Economía Aplicada Cuantitativa, Facultad de Ciencias Económicas y Empresariales, UNED, 28040 Madrid, Spain

(Received 29 January 1996; accepted for publication 10 April 1996)

We present a geometric framework for non-holonomic Lagrangian systems in terms of distributions on the configuration manifold. If the constrained system is regular, an almost product structure on the phase space of velocities is constructed such that the constrained dynamics is obtained by projecting the free dynamics. If the constrained system is singular, we develop a constraint algorithm which is very similar to that developed by Dirac and Bergmann, and later globalized by Gotay and Nester. Special attention to the case of constrained systems given by connections is paid. In particular, we extend the results of Koiller for Čaplygin systems. An application to the so-called non-holonomic geometry is given. © *1996 American Institute of Physics.* [S0022-2488(96)02407-3]

I. INTRODUCTION

A non-holonomic Lagrangian system consists of a regular Lagrangian $L(q^A,\dot{q}^A)$ defined on the phase space of velocities TQ of a configuration manifold Q with local coordinates $(q^A), 1 \leqslant A \leqslant n = \dim Q$, subjected to constraints defined by m local functions $\phi_i(q^A,\dot{q}^A)$. That means that the only allowable velocities are those verifying that $\phi_i = 0$. We only consider the case of linear constraints, say those of the form $\phi_i(q^A,\dot{q}^A) = (\mu_i)_A(q)\dot{q}^A$. By applying a suitable Hamilton's principle, we arrive to the constrained Euler–Lagrange equations,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^A}\right) - \frac{\partial L}{\partial q^A} = -\lambda^i(\mu_i)_A,$$

where λ^i, $1 \leqslant i \leqslant m$, are some Lagrange multipliers to be determined (see, for instance, Valcovici,[1] Pars,[2] Neimark and Fufaev,[3] Vershik and Faddeev,[4] Saletan and Cromer,[5] Rumiantsev,[6] Pironneau,[7] Vershik and Gershkovich,[8] Massa and Pagani[9,10]). In some of them, a more general type of constraints was discussed. We notice that Hamilton's principle in the non-holonomic framework is not a variational principle. We remit to the excellent book by Rosenberg[11] for a detailed discussion on that subject.

In the last years, there is an increasing interest in non-holonomic mechanics, and other approaches from a geometrical point of view have appeared: Weber,[12] Pitanga,[13,14] Marle,[15] Massa and Pagani,[9,10] Bates and Śniatycki,[16] Giachetta,[17] Koiller,[18] Cariñena and Rañada,[19] Rañada,[20] Dazord,[21] Cariñena and Rañada,[22] Sarlet, Cantrijn and Saunders,[23,24] Sarlet,[25,26] de León and M. de Diego.[27–31]

Our approach is a globalization of the one by Cariñena and Rañada.[19] In order to globalize their picture, we will consider a distribution D of codimension m defined on Q. The constraints

a)Electronic mail: mdeleon@pinar1.csic.es
b)Electronic mail: ceedd05@cc.csic.es

imply that the motion is only allowable for some values of velocities, those belonging to the distribution D. The constrained motion equations can be written out by modifying the motion equations for the associated free Lagrangian system as follows:

$$(i_X\omega_L - dE_L) \in (D^v)^0, \quad X \in D^c, \tag{1}$$

along the points of D, where ω_L is the symplectic Poincaré–Cartan two-form, E_L is the energy associated with L, and D^v and D^c are the lifts of D to TQ. Notice that we do not need to invoke Lagrange multipliers. This approach is the dual version (i.e., in terms of distributions) of the one by Cartan using exterior systems.

Under some regularity hypothesis we construct an almost product structure $(\mathscr{P},\mathscr{Q})$ on TQ along the linear submanifold D such that the dynamics are obtained by projecting the Euler-Lagrange vector field ξ_L which solves the motion equations of the free problem,

$$i_{\xi_L}\omega_L = dE_L.$$

That is, the solutions of the constrained dynamics are just the solutions of the second order differential equation $\xi = \mathscr{P}(\xi_L)$ (Section II).

If the constrained system is not regular, we develop in Section III a constraint algorithm which is remarkably similar to that developed by Dirac and Bergmann for singular Lagrangian systems.[32–34] We obtain the local and global aspects of the constraint algorithm. By the way, we introduce the notion of first and second class constraints in this framework.

In Section IV we consider a very important kind of constrained systems, those called generalized Čaplygin systems. A generalized Čaplygin system consists of a Lagrangian function $L:TQ \to \mathbb{R}$ and a connection Γ in a fibration $\rho:Q \to M$ such that L is invariant by the horizontal lift operation. The particular case when $\rho:Q \to M$ is a principal bundle with structure group G, L is G-invariant and Γ is a principal connection, i.e., the horizontal subspaces are G-invariant, was considered by Koiller.[18] We extend the results by Koiller, and prove that there exists a well-defined Lagrangian function $L^*:TM \to \mathbb{R}$, such that the generalized Čaplygin system (L,Γ) is equivalent to a non-conservative system on TM with Lagrangian function L^* and external force α. Here, α is an one-form on TM related with the curvature of the connection Γ. Roughly speaking, the curvature is just the force of constraint. Several examples are studied.

Finally, in Section V, we apply our procedure to give a new insight to an old problem in the so-called non-holonomic geometry. Let Q be a Riemannian manifold with Riemannian metric g and Levi-Civita connection ∇ and suppose that a distribution D on Q is given. The goal is to obtain a new linear connection ∇^* on Q such that the geodesics of ∇^* are the extremals of the variational problem subjected to these linear constraints (see Synge,[35] Vranceanu,[36] Neimark and Fufaev[3] and the references therein). We define a connection Γ^* along D by using our procedure and the relations between non-homogeneous connections and second order differential equations on TQ obtained by Grifone.[37] If the constraints are holonomic, Γ^* induces a linear connection in the vector bundle $D \to Q$.

II. NON-HOLONOMIC LAGRANGIAN SYSTEMS

Let $L:TQ \to \mathbb{R}$ be a Lagrangian function defined on the phase space of velocities TQ of a n-dimensional configuration manifold Q. Denote by (q^A, v^A) the fibered coordinates on TQ. (Sometimes we will use the notation $v^A = \dot{q}^A$.)

The motion equations for L can be derived by a variational procedure. In fact, the extremals of the action,

$$\int L(q^A, \dot{q}^A)\,dt,$$

where $\dot{q}^A = dq^A/dt$ are just the solutions of the Euler–Lagrange equations,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial v^A}\right) - \frac{\partial L}{\partial q^A} = 0, \quad v^A = \frac{dq^A}{dt}. \tag{2}$$

Alternatively, there exists a symplectic formulation. Indeed, let $\omega_L = -d\alpha_L$ be the Poincaré-Cartan 2-form, where $\alpha_L = J^*(dL)$. Here, J is the canonical almost tangent structure on TQ. That is, J is a tensor field of type (1,1) on TQ such that $J^2 = 0$ and rank $J = n$; J is locally defined by

$$J = \frac{\partial}{\partial v^A} \otimes dq^A.$$

The energy associated with L is defined by $E_L = CL - L$, where $C = v^A(\partial/\partial v^A)$ is the Liouville vector field on TQ. That is, C is the infinitesimal generator of the dilations along the fibers.

The global motion equation for the free problem is (see Ref. 38)

$$i_X \omega_L = dE_L. \tag{3}$$

We say that the Lagrangian L is regular if the Hessian matrix $(\partial^2 L/\partial v^A \partial v^B)$ is non-singular. In such a case, the form ω_L is symplectic and, thus, (3) has a unique solution ξ_L (the Euler–Lagrange vector field) which is a second order differential equation (SODE for short). Further, the solutions of ξ_L coincide with the solutions of the Euler–Lagrange equations. More precisely, the projections onto Q of the integral curves of ξ_L are the extremals for L.

Now, we suppose that a family of linear constraints is given. In the local picture, L is subjected to constraints defined by m local functions of the form $\phi_i(q^A, v^A) = (\mu_i)_A(q) v^A$. That means that the only allowable velocities are those verifying that $\phi_i = 0$.

The purpose of this paper is to give a global picture of Lagrangian systems subjected to linear constraints.

Definition II.1: A non-holonomic Lagrangian system is given by the following data:
(i) A regular Lagrangian $L: TQ \to \mathbb{R}$*;*
(ii) An $(n-m)$*-dimensional distribution* D *on the* n*-dimensional configuration manifold* Q.
The constraints are said to be holonomic if D *is involutive.*

This means that the only allowable velocities are the tangent vectors belonging to D, i.e., the motion is constrained to the submanifold D. Notice that D can be viewed as a vector subbundle of $\tau_Q: TQ \to Q$, and, so, D is a submanifold of TQ.

Define two distributions D^c and D^v on TQ as follows. Suppose that $\{\mu_i; 1 \leq i \leq m\}$ is a local basis of 1-forms of the annihilator D^0 of D. Then D^c and D^v are, respectively, defined by

$$(D^c)^0 = \langle \mu_i^v, \mu_i^c \rangle, \quad (D^v)^0 = \langle \mu_i^v \rangle, \tag{4}$$

where μ_i^v (resp., μ_i^c) denotes the vertical (resp., complete) lift of the 1-form μ_i to TQ (see Refs. 38,39).

If $\{\overline{\mu_i}\}$ is another local basis of D we have

$$\overline{\mu_i} = \Lambda_i^j \mu_j,$$

where (Λ_i^j) is a regular matrix defined on the overlapping of the two local neighborhoods. Since

$$\overline{\mu_i^c} = (\Lambda_i^j)^c \mu_j^v + (\Lambda_i^j)^v \mu_j^c, \quad \overline{\mu_i^v} = (\Lambda_i^j)^v \mu_j^v, \tag{5}$$

we deduce that D^c and D^v are well-defined. Here f^v (resp., f^c) denotes the vertical (resp., complete) lift of a function f on Q to TQ (see Refs. 39, 38).

Since the allowable velocities have to belong to D, we deduce that the motion equations for the constrained mechanical systems would be

$$(i_X\omega_L - dE_L) \in (D^v)^0, \quad X \in D^c, \tag{6}$$

along the points of D.

In fact, suppose that $\mu_i = (\mu_i)_A dq^A$. Hence we have

$$\mu_i^v = (\mu_i)_A dq^A,$$
$$\mu_i^c = (\mu_i)_A^c dq^A + (\mu_i)_A dv^A,$$
$$\hat{\mu}_i = (\mu_i)_A v^A,$$

where $\hat{\mu}_i$ is the function on TQ defined by $\hat{\mu}_i(q^A, v^A) = \mu_i(v^A(\partial/\partial q^A))$.

Notice that there are many solutions of the first equation in (6), since ω_L is symplectic. Moreover, every solution of (6) is a SODE. In fact, let μ_i be a local basis of D. Then, (6) can be locally written as follows:

$$i_X\omega_L - dE_L = \lambda^i \mu_i^v, \quad \mu_i^v(X) = 0, \quad \mu_i^c(X) = 0, \tag{7}$$

for some Lagrange multipliers λ^i to be determined. If we apply i_J to the first equation in (7), we have

$$i_J i_X \omega_L - i_J(dE_L) = 0,$$

which implies $i_{JX}\omega_L = i_C\omega_L$ and, then, $JX = C$. Here, i_J denotes the derivation of type i_* in the sense of Frölicher–Nijenhuis associated with J, that is, i_J is completely defined by the formulas $i_J f = 0$ and $i_J(df) = J^*(df)$ for any function f on TQ.[38] Therefore, we deduce that the local expression of X is

$$X = v^A \frac{\partial}{\partial q^A} + X^A \frac{\partial}{\partial v^A}.$$

Thus, the solutions of X satisfy the following constrained Euler–Lagrange equations:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial v^A}\right) - \frac{\partial L}{\partial q^A} = -\lambda^i (\mu_i)_A, \quad v^A = \frac{dq^A}{dt}.$$

By the way, notice that $\mu_i^v(X) = \hat{\mu}_i$ and $\mu_i^c(X) = X(\hat{\mu}_i)$. Then the second set of equations in (7) defines the submanifold D, and the third one means that X has to be tangent to D. Given the symplectic form ω_L we have the associated musical isomorphisms,

$$\flat_L : T(TQ) \to T^*(TQ)$$

and

$$\# : T^*(TQ) \to T(TQ),$$

where for any $X \in T(TQ)$, $\flat_L(X) = i_X\omega_L$, and for any $\alpha \in T^*(TQ)$, $\#\alpha = X_\alpha$ is the unique tangent vector to TQ such that $b_L(X_\alpha) = \alpha$. From $(D^v)^0$ and by using the isomorphism $\#$, we obtain a distribution S on TQ. A vector $v \in T_x TQ$ belongs to S_x if $i_v\omega_L(x) \in (D^v)_x^0$. It is clear that $\dim S = m$.

Let Z_i be the local vector field defined by

$$i_{Z_i}\omega_L=\mu_i^v, \quad 1\leqslant i\leqslant m.$$

A direct computation shows that

$$Z_i=-W^{AB}(\mu_i)_B\frac{\partial}{\partial v^A},$$

where (W^{AB}) is the inverse matrix of the Hessian matrix $(W_{AB}=\partial^2L/\partial v^A\partial v^B)$. Therefore, Z_i is vertical.

We have that S is locally generated by the vector fields $\{Z_i,\ 1\leqslant i\leqslant m\}$.

Definition II.2: The constrained system will be called regular if

$$T_xD\cap S_x=0, \quad \textit{for any} \ \ x\in D.$$

The meaning of the regularity of the constrained system will become clear in a while. Notice that the regularity of the constrained system is closely related with the nature of the Lagrangian function.

Suppose that the constrained system is regular. Since for any $x\in D$, we have $\dim S_x=m$, and $\dim T_xD=2n-m$, we obtain that

$$T_xTQ=T_xD\oplus S_x, \quad \forall x\in D.$$

Thus, each tangent vector $v\in T_xTQ$ splits in a unique way as $v=v_1+v_2$, where $v_1\in T_xD$ and $v_2\in S_x$. Then, we can construct two complementary projectors $\mathscr{P}$ and $\mathscr{Q}$ as follows: $\mathscr{P}(v)=v_1$, and $\mathscr{Q}(v)=v_2$. In fact, $(\mathscr{P},\mathscr{Q})$ is a well-defined almost product structure on TQ along the points of D.

Take now the codistribution $\langle dE_L\rangle \oplus (D^v)^0$. By using the isomorphism $\#$, we obtain a distribution S_L on TQ locally generated by $\{\xi_L, Z_1,\cdots,Z_m\}$. That is,

$$S_L=\#(\langle dE_L\rangle\oplus(D^v)^0).$$

Then, $\dim(T_xD\cap(S_L)_x)=1$, $\forall x \in D$. Moreover, there exists a unique generator ξ of the distribution $TD\cap S_L$ along the points of D such that $(J\xi=C)_{/D}$.

The vector field $\xi \in \mathscr{X}(D)$ is the solution of the Lagrangian system subjected to constraints given by a distribution D. This vector field ξ is precisely $P(\xi_{L/D})$, the projection of the Euler-Lagrange vector field of the free Lagrangian system. In fact, along the points of D, we have

$$\begin{aligned} i_{\mathscr{P}(\xi_L)}\omega_L-dE_L&=i_{\xi_L-\mathscr{Q}(\xi_L)}\omega_L-dE_L\\ &=-i_{\mathscr{Q}(\xi_L)}\omega_L\in(D^v)^0. \end{aligned}$$

Moreover, we deduce that $\mathscr{P}(\xi_L)(x)\in T_xD$ or, equivalently, $\mathscr{P}(\xi_L)\in D^c$.

In order to perform an explicit computation of the vector field ξ, we proceed as follows. Take a local basis $\{\mu_i, 1\leqslant i\leqslant m\}$ of D and define $\mathscr{C}_{ij}=Z_i(\hat{\mu}_j)$. We deduce that

$$\mathscr{C}_{ij}=-W^{AB}(\mu_i)_A(\mu_j)_B. \tag{8}$$

Proposition II.3: The constrained system is regular iff *the matrices* $(\mathscr{C}_{ij})$ *are non-singular on* D.

Proof: Suppose that the constrained system is regular. Take an arbitrary linear combination of columns of $\mathscr{C}$ at some point $x\in D$ such that

$$\sum_{i=1}^{m} \lambda^i Z_i(x)(\hat{\mu}_j)=0.$$

Thus, $\Sigma\lambda^i Z_i(x)\in T_xD$ which implies that $\Sigma\lambda^i Z_i(x)=0$ and hence $\lambda^1=\lambda^2=\cdots\lambda^m=0$.

Conversely, suppose $\mathscr{C}$ ia non-singular and take $X\in S_x\cap T_xD$. Thus, $X=\Sigma\lambda^i Z_i(x)$ and $X(\hat{\mu}_j)=0$, $\forall j$, $1\leqslant j\leqslant m$ which implies that $\Sigma\lambda^i Z_i(\hat{\mu}_j)=0$. Therefore we deduce that $\lambda^1=\cdots=\lambda^m=0$ and $X=0$.

■

Thus, if D is regular, we obtain an explicit expression for the projector $\mathscr{Q}$:

$$\mathscr{Q}=\mathscr{C}^{ij}Z_j\otimes d\hat{\mu}_i,$$

where $(\mathscr{C}^{ij})$ denotes the inverse matrix of $(\mathscr{C}_{ij})$.

Then, we get

$$\xi=\mathscr{P}(\xi_L)=\xi_L-\mathscr{C}^{ij}\xi_L(\hat{\mu}_i)Z_j.$$

Proposition II.4: If the Hessian matrix

$$\left(\frac{\partial^2 L}{\partial v^A\partial v^B}\right),$$

is positive or negative definite at each point $x\in D$, then the constrained system is regular.

Proof: The result follows from (8) (see also Cariñena and Rañada[19]).

■

Remark II.5: Proposition II.4 clarifies the usual assumption on the positive or negative character of the Hessian matrix of L. It is nothing but that a sufficient condition to ensure the regularity of the constrained system. Of course, if the Lagrangian L is natural, that is, $L=T-V$, where T is the kinetic energy of a Riemannian metric g on Q and V is a potential energy, then the constrained system would be regular.

Remark II.6: From the regularity of the matrices $\mathscr{C}$, we deduce that $(\mathscr{P},\mathscr{Q})$ may be extended (in many ways) to an open neighborhood of D. Consequently, $\mathscr{P}(\xi_L)$ may also be extended to an open neighborhood of D (see Ref. 27 for more details).

By using the almost product structure $(\mathscr{P},\mathscr{Q})$ and the musical isomorphisms, we can construct the following linear mapping $\overline{\mathscr{Q}}_x:T_x^*(TQ)\to T_x^*(TQ)$:

$$\overline{\mathscr{Q}}_x(\alpha_x)=\flat_L(\mathscr{Q}_x(\#(\alpha_x))),\quad \forall\alpha_x\in T_x^*(TQ),\quad x\in D.$$

Since $\overline{\mathscr{Q}}_x^2=\mathrm{id}$ and $Im\overline{\mathscr{Q}}_x=(D^v)_x^o$, we obtain the following splitting:

$$T_x^*(TQ)=(D^v)_x^o\oplus\overline{S_x},\quad \forall x\in D,$$

where $\overline{S_x}=Im\overline{\mathscr{P}}_x$, $\overline{\mathscr{P}}_x=\mathrm{id}-\overline{\mathscr{Q}}_x$ being the complementary projector. In fact, $\overline{\mathscr{P}}$ and $\overline{\mathscr{Q}}$ may be interpreted as tensor fields of type (1,1) on TQ defined along D.

Notice that $\overline{S}$ is the annihilator of the distribution along D locally generated by the vector fields $\{X_{\hat{\mu}_i},1\leqslant i\leqslant m\}$, where $X_{\hat{\mu}_i}$ is the Hamiltonian vector field of the function $\hat{\mu}_i$ with respect to the symplectic form ω_L.

The following result tells us that one could add the constraint forces to the energy to obtain a global force acting on the system.

Theorem II.7: *The solution of the constrained dynamics is the unique vector field ξ on D such that*

$$(i_X\omega_L=\overline{\mathscr{P}}(dE_L))_{/D}.$$

Proof: By the construction of $\overline{\mathscr{Q}}$ we have that

$$\overline{\mathscr{Q}}(dE_L)=i_{\mathscr{Q}(\xi_L)}\omega_L.$$

On the other hand, $\xi=\mathscr{P}(\xi_L)$ is the solution of the constrained dynamics. Thus, we deduce that

$$i_{\mathscr{P}(\xi_L)}\omega_L=i_{\xi_L}\omega_L-i_{\mathscr{Q}(\xi_L)}\omega_L=dE_L-\overline{\mathscr{Q}}(dE_L)=\overline{\mathscr{P}}(dE_L).$$

Since ω_L is symplectic, we conclude that the solution of the equation

$$(i_X\omega_L=\overline{\mathscr{P}}(dE_L))_{/D}$$

is unique. ■

A direct computation shows that the local expression of $\overline{\mathscr{Q}}$ is

$$\overline{\mathscr{Q}}=-\mathscr{C}^{ij}X_{\hat{\mu}_i}\otimes\mu_j^v.$$

Therefore, we obtain that

$$\overline{\mathscr{P}}(dE_L)=dE_L-\mathscr{C}^{ij}\xi_L(\hat{\mu}_i)\mu_j^v.$$

The following lemmas will be used in Section IV.

Lemma II.8: Given a regular constrained system with Lagrangian function L and linear constraints D, the vector field ξ solving the constrained dynamics satisfies

$$\mathscr{L}_\xi\alpha_L=dL-L_{\mathscr{Q}(\xi_L)}\alpha_L,$$

where $\mathscr{L}_\xi$ denotes the Lie derivative with respect to ξ.

Proof: It follows since $\xi=\mathscr{P}(\xi_L)=\xi_L-\mathscr{Q}(\xi_L)$ and $\mathscr{L}_{\xi_L}\alpha_L=dL$. ■

Lemma II.9: Under the same hypothesis as in Lemma II.8, we have

$$\mathscr{L}_{\mathscr{Q}(\xi_L)}\alpha_L\in(D^v)^0.$$

Proof: Since $\mathscr{Q}(\xi_L)=\Sigma_{j=1}^m\Lambda^jZ_j$, with $\Lambda^j=\mathscr{C}^{ij}\xi_L(\hat{\mu}_i)$, we deduce that

$$\begin{aligned}\mathscr{L}_{\mathscr{Q}(\xi_L)}\alpha_L&=\mathscr{L}_{\sum_{j=1}^m\Lambda^jZ_j}\alpha_L\\&=i_{\sum_{j=1}^m\Lambda^jZ_j}d\alpha_L+d(i_{\sum_{j=1}^m\Lambda^jZ_j}\alpha_L)\\&=-i_{\sum_{j=1}^m\Lambda^jZ_j}\omega_L=-\sum_{j=1}^m\Lambda^j\mu_j^v,\end{aligned}$$

since the vector fields Z_j are vertical and α_L is semibasic. ■

III. THE SINGULAR CASE

In this section we shall describe what happens if the given constrained system is not regular, or, in other words, $T_xD\cap S_x=0$, for some points x in D. Notice that this fact is equivalent to the non-regularity of the local matrices $(\mathscr{C}_{ij})$.

For any point $x\in D$, we have the obvious inclusion

$$T_xD\cap S_x\subset T_xD\cap (S_L)_x .$$

In the regular case, this inclusion is strict, and the jump of dimension is just 1. This jump allows us to obtain the dynamics by taking a basis ξ of $(TD\cap S_L)_{/D}$ normalized in order to get $J\xi=C$. The above remarks illuminate the way to proceed in the singular case.

Define a submanifold D_2 of TQ as follows:

$$D_2=\{x\in D/T_xD\cap S_x\subsetneqq T_xD\cap (S_L)_x\}.$$

This implies that for any point $x\in D_2$, there exists some tangent vector

$$X=\xi_L(x)+\lambda^iZ_i(x)\in T_xD\cap (S_L)_x ,$$

such that $X\notin S_x$. Thus, X is a solution of the constrained equation, but, in general, it is not necessarily tangent to D_2.

Therefore, we define a new submanifold D_3 of D_2 as follows:

$$D_3=\{x\in D_2/T_xD_2\cap S_x\subsetneqq T_xD_2\cap (S_L)_x\}.$$

Proceeding further, we obtain a sequence of constraint submanifolds,

$$\cdots\rightarrow D_k\rightarrow\cdots\rightarrow D_2\rightarrow D_1=D.$$

$D=D_1$ will be called the primary constraint submanifold, D_2 the secondary constraint submanifold and so on.

As in the Gotay and Nester algorithm for singular Lagrangian systems[33,34] we also have three possibilities

(i) There exists an integer $k\geqslant 1$ such that $D_k=\emptyset$. This means that the equations (6) are not consistent.

(ii) There exists an integer $k\geqslant 1$ such that $D_k=\emptyset$ but $\dim D_k=0$. In this case, there are no dynamics. D_k consists in isolated points and the solution of the constrained dynamics is $X=0$.

(iii) There exists an integer $k\geqslant 1$ such that $D_{k+1}=D_k$ and $\dim D_k>0$. In this case the algorithm stabilizes at the final constraint submanifold $D_f=D_k$. So, there exists at least a vector field ξ on D_f satisfying the SODE condition $((J\xi=C)_{/D_f})$ and such that

$$i_\xi\omega_L-dE_L\in (D^v)^0.$$

Assume that the algorithm ends at some final constraint submanifold D_f. Thus, we have

$$T_xD_f\cap S_x\subsetneqq T_xD_f\cap (S_L)_x ,\quad \forall x\in D_f .$$

We will suppose that the distribution $TD_f\cap S$ along D has constant dimension, say r, that is,

$$\dim(T_xD_f\cap S_x)=r,\quad \text{for all }\ x\in D_f .$$

Lemma III.1: We have that $\xi_L(x)\in T_xD_f+S_x$ for any $x\in D_f$.

Proof: Since $T_xD_f\cap S_x\subsetneqq T_xD_f\cap (S_L)_x$ for any $x\in D_f$, we deduce that there exists a tangent vector $X\in T_xD_f\cap (S_L)_x$ such that $X\notin T_xD_f\cap S_x$. Thus, $X=\xi_L(x)+\lambda^iZ_i(x)$ which implies that

$\xi_L(x)=X-\lambda^i Z_i \in T_xD_f+S_x$.

■

In order to determine the dynamics, we split S_x as a direct sum of two subspaces, say

$$S_x=\check{S}_x \oplus (T_xD_f \cap S_x). \tag{9}$$

Obviously, there are many choices for a complementary subspace of $T_xD_f \cap S_x$ into S_x. From (9) we deduce that $\check{S}_x \cap T_xD_f=0$, and we can then split $T_x(TQ)$ as follows:

$$T_x(TQ)=\check{S}_x \oplus T_xD_f \oplus M_x, \quad x \in D_f,$$

where M_x is a suitable complementary subspace. Take the corresponding three complementary projectors:

$$\mathscr{Q}_x : T_x(TQ) \to \check{S}_x,$$

$$(\mathscr{P}_1)_x : T_x(TQ) \to T_xD_f,$$

$$(\mathscr{P}_2)_x : T_x(TQ) \to M_x.$$

Consider the projector $(\mathscr{P})_x=(\mathscr{P}_1)_x+(\mathscr{P}_2)_x$. Hence, we have (along the points of D_f)

$$\begin{aligned} i_{\mathscr{P}_x(\xi_L(x))}\omega_L(x)-(dE_L)_x &= i_{\xi_L(x)-\mathscr{Q}_x(\xi_L(x))}\omega_L(x)-i_{\xi_L(x)}\omega_L(x) \\ &= -i_{\mathscr{Q}_x(\xi_L(x))}\omega_L(x) \in (D^v)^0_x. \end{aligned}$$

Moreover, for any $x \in D_f$, we deduce that

$$\mathscr{P}_x(\xi_L(x))=(\mathscr{P}_1)_x(\xi_L(x)) \in T_xD_f,$$

since $\xi_L(x) \in T_xD_f+S_x$ by Lemma III.1.

A differentiable choice of both distributions $\check{S}$ and M allows us to construct an almost product structure $(\mathscr{P}_1,\mathscr{P}_2,\mathscr{Q})$ [or $(\mathscr{P},\mathscr{Q})$, where $\mathscr{P}=\mathscr{P}_1+\mathscr{P}_2$] along D_f such that $\mathscr{P}(\xi_{L/D_f})$ is a solution of the constrained dynamics. Notice that a general solution is of the form

$$\mathscr{P}(\xi_{L/D_f})+TD_f \cap S.$$

We have chosen complementary distributions $\check{S}$ and M in order to obtain the dynamics. Notice that it is possible to realize both decompositions, say $S=\check{S} \oplus (T_xD_f \cap S)$ and $T(TQ)=\check{S} \oplus TD_f \oplus M$. In fact, take a local basis $\{\mu_i\}$ of D^0 and denote by ϕ_I the constraint functions which define D_f, where $1 \leqslant I \leqslant 2n-\dim D_f$. Notice that $\phi_I=\hat{\mu}_I$, for $1 \leqslant I \leqslant m$. We have assumed that $TD_f \cap S$ has constant rank r. Thus, the matrix $(\mathscr{C}_{iJ})=(Z_i(\hat{\phi}_J))$ has also constant rank $m-r$. Indeed, take a local basis $Y_1,\ldots,Y_r$ of $TD_f \cap S$ such that $Y_a=\mathscr{A}^i_a Z_i$. Since Y_a is tangent to D_f, we get

$$\mathscr{A}^i_a Z_i(\phi_J)=0, \quad \text{for all } J.$$

But this implies that rank $\mathscr{C}=m-r$. The converse is proved by reversing the argument.

Assume that the submatrix $\mathscr{C}'=(\mathscr{C}_{I'J'})$, $(1 \leqslant I',J' \leqslant m-r)$ is regular. In that case, we define a projector $\mathscr{Q}$ by putting

$$\mathscr{Q}=\mathscr{C}^{I'J'}Z_{J'} \otimes d\phi_{I'},$$

where $(\mathscr{C}^{I'J'})$ is the inverse matrix of $(\mathscr{C}_{I'J'})$. Notice that $\check{S}=\langle Z_{I'}\rangle$. If we put $\mathscr{P}=\mathrm{id}-\mathscr{Q}$ we obtain an almost product structure $(\mathscr{P},\mathscr{Q})$ along D_f. The decomposition $\mathscr{P}=\mathscr{P}_1+\mathscr{P}_2$ is obtained by choosing a complementary M of $\check{S}\oplus TD_f$. This choice corresponds to the ambiguity in the determination of the other Lagrange multipliers. Indeed, if we compute $\mathscr{P}(\xi_L)$ we obtain

$$\mathscr{P}(\xi_L)=\xi_L-\mathscr{C}^{I'J'}\xi_L(\phi_{I'})Z_{J'},$$

and a general solution is of the form

$$\mathscr{P}(\xi_L)+Y,$$

where $Y\in TD_f\cap S$. So, the only Lagrange multipliers determined are just the components of the $Z_{J'}$'s.

Remark III.2: A solution of (6) has the general form $X=\xi_L+\lambda^i Z_i$, where Z_i are the symplectic gradients of the 1-forms μ_i^v. The tangency condition may now be written as

$$\xi_L(\hat{\mu}_j)+\lambda^i Z_i(\hat{\mu}_j)=0,\quad 1\leq j\leq m, \tag{10}$$

or, equivalently,

$$\xi_L(\hat{\mu}_j)+\lambda^i\mathscr{C}_{ij}=0,\quad 1\leq j\leq m. \tag{11}$$

If the matrix $\mathscr{C}_{ij}$ is regular, the system of equations (11) have a solution which is obtained by the well-known Crame rule, or, in a more sophisticated way, by constructing the local almost product structure $(\mathscr{P},\mathscr{Q})$.

If the constrained system is singular, Equation (10) can be analyzed as in the Dirac-Bergmann algorithm.[32] In fact,

$$\xi_L(\hat{\mu}_j)+\lambda^i\mathscr{C}_{ij}=0,\quad 1\leq j\leq m,$$

is a system of m equations with m unknowns, the Lagrange multipliers. The system is consistent if the ranks of the matrices $(\mathscr{C}_{ij})$ and $(\mathscr{C}_{ij};-\xi_L(\hat{\mu}_j))$ coincide. (Of course, they are equal if the constrained system is regular.) Therefore, we select the points where the ranks coincide. Denote by $\overline{D_2}$ the collection of all these points. At the points in $\overline{D_2}$ there are solutions, but they are not necessarily tangent to $\overline{D_2}$. By the way, new constraints may appear. In fact, notice that, if the matrix $(\mathscr{C}_{ij})$ has rank, say r, then the matrix $(\mathscr{C}_{ij};-\xi_L(\hat{\mu}_j))$ has rank greater or equal to r. Suppose that $\mathscr{M}$ is a submatrix of $(\mathscr{C}_{ij})$ of rank r. The determinants of the submatrices of $(\mathscr{C}_{ij};-\xi_L(\hat{\mu}_j))$ obtained from $\mathscr{M}$ by adding elements of the column $\xi_L(\hat{\mu}_j)$ are the new possible constraints. These secondary constraints ϕ_α have to be added to the motion equations which become

$$\begin{aligned}\xi_L(\hat{\mu}_j)+\lambda^i Z_i(\hat{\mu}_j)=0,\\ \xi_L(\phi_\alpha)+\lambda^i Z_i(\phi_\alpha)=0.\end{aligned} \tag{12}$$

The procedure is now repeated, and we obtain a sequence of submanifolds

$$\cdots\rightarrow\overline{D_k}\rightarrow\cdots\rightarrow\overline{D_2}\rightarrow D,$$

which are just the same that the ones previously obtained. More precisely, $\overline{D_k}$ is the intersection of D_k with the tangent bundle of the open neighborhood where the local basis μ_i is defined.

Remark III.3: We started with linear constraints, or, in the present terminology, the primary constraints are linear. However, the secondary constraints are not necessarily linear.

If we denote by $\{,\}_L$ the Poisson bracket on TQ defined from the symplectic form ω_L, we have $\{f,E_L\}_L=\xi_L(f)$, for any function f on TQ. Thus, (10) can be written as follows:

$$\{\hat{\mu}_j,E_L\}_L+\lambda^i Z_i(\hat{\mu}_j)=0,\quad 1\leqslant j\leqslant m.$$

As in the Dirac–Bergmann approach,[32] we can distinguish two different classes of constraints. A constraint f will be called first class if $\{f,\Psi\}_L\sim 0$, that is, $\{f,\Psi\}_L$ vanishes on the final constraint submanifold D_f. Otherwise, f will be called a second class constraint. We deduce that the Hamiltonian vector fields corresponding to first class constraints are tangent to D_f, and the Hamiltonian vector fields corresponding to second class constraints are transversal to D_f.

Example III.4: Consider the following Lagrangian function L defined on $T\mathbb{R}^3$ by

$$L=\frac{1}{2}((v^1)^2+(v^2)^2-(v^3)^2+(q^1)^2),$$

subjected to linear constraints given by a distribution D on $\mathbb{R}^3$ whose annihilator is

$$D^0=\langle dq^1+dq^3\rangle.$$

Here (q^1,q^2,q^3) denote the standard coordinates on $\mathbb{R}^3$, and $(q^1,q^2,q^3,v^1,v^2,v^3)$ the induced ones on $T\mathbb{R}^3$. Thus, the submanifold $D\subset T\mathbb{R}^3$ consists in those points in $T\mathbb{R}^3$ such that $v^1+v^3=0$.

The distribution S is generated by the vector field

$$Z=-\frac{\partial}{\partial v^1}+\frac{\partial}{\partial v^3}.$$

We have

$$S_x\cap T_xD=S_x,$$

for any $x\in D$. Therefore, the constrained system is singular.

Applying the constraint algorithm, we get

$$D_2=\{x\in D/T_xD\cap S_x\subsetneqq T_xD\cap(S_L)_x\},$$

so that

$$D_2=\{(q^A,v^A)\in T\mathbb{R}^3/v^1+v^3=0,\ q^1=0\}.$$

We proceed further, and obtain

$$\begin{aligned}D_3&=\{x\in D/T_xD_2\cap S_x\subsetneqq T_xD_2\cap(S_L)_x\},\\ &=\{(q^A,v^A)\in T\mathbb{R}^3/v^3=0,\ q^1=0,\ v^1=0\}.\end{aligned}$$

Now, since

$$T_xD_3\cap S_x\subsetneqq T_xD_3\cap(S_L)_x,$$

for any $x\in D_3$, we deduce that D_3 is the final constraint submanifold.

The dynamics is given by the vector field $(\xi_L+\lambda Z)_{/D_3}$, for any function λ on D_3.

IV. CONSTRAINTS DEFINED BY CONNECTIONS. GENERALIZED ČAPLYGIN SYSTEMS

One of the most appealing instances of non-holonomic Lagrangian systems are those given by the existence of a connection.

Suppose that Q is a fibered manifold over a manifold M, say, $\rho:Q\to M$ is a surjective submersion. Assume that a connection Γ in $\rho:Q\to M$ is given, such that the allowable motions of a Lagrangian function $L:TQ\to\mathbb{R}$ have to be horizontal curves with respect to that connection. In other words, the allowable velocities are horizontal tangent vectors. Thus, D is just the horizontal distribution H such that

$$TQ=H\oplus V\rho.$$

Let us recall that Γ may be considered as a tensor field of type (1,1) on Q such that $\Gamma^2=$ id and the eigenspaces corresponding to the eigenvalue -1 are just the vertical subspaces. Take fibered coordinates $(q^A)=(q^a,q^i)$, $1\leqslant a\leqslant n-m$, $1\leqslant i\leqslant m$, $n=\dim Q$. The horizontal distribution is locally spanned by the local vector fields

$$H_a=\left(\frac{\partial}{\partial q^a}\right)^H=\frac{\partial}{\partial q^a}-\Gamma_a^i(q^A)\frac{\partial}{\partial q^i},$$

where Y^H stands for the horizontal lift to Q of a vector field Y on M, and $\Gamma_a^i=\Gamma_a^i(q^b,q^j)$ are the Christoffel components of Γ. Thus, we obtain a local basis of vector fields on Q,

$$\left\{H_a,V_i=\frac{\partial}{\partial q^i}\right\}.$$

Its dual basis of 1-forms is

$$\{\eta_a=dq^a,\eta_i=\Gamma_a^i dq^a+dq^i\}.$$

We deduce that H^0 is locally spanned by the 1-forms $\{\eta_i\}$.

Define the curvature of Γ as the tensor field of type (1,2) on Q given by

$$R=\frac{1}{2}[\mathbf{h},\mathbf{h}],$$

where $\mathbf{h}=(1/2)(\text{ id}+\Gamma)$ is the horizontal projector associated with Γ, and $[\mathbf{h},\mathbf{h}]$ is its Nijenhuis tensor (see Ref. 38). Thus,

$$R(\mathbf{h}(u_1),\mathbf{h}(u_2))=v([\mathbf{h}(u_1),\mathbf{h}(u_2)]),$$

$$R(\mathbf{h}(u_1),\mathbf{v}(u_2))=0,$$

$$R(\mathbf{v}(u_1),\mathbf{v}(u_2))=0,$$

for any $u_1,u_2\in T_xQ$, where $\mathbf{v}=\text{ id}-\mathbf{h}$ is the complementary vertical projector. Since

$$\mathbf{h}\left(\frac{\partial}{\partial q^a}\right)=\frac{\partial}{\partial q^a}-\Gamma_a^i\frac{\partial}{\partial q^i},$$

$$\mathbf{h}\left(\frac{\partial}{\partial q^i}\right)=0,$$

we obtain

$$R\left(\frac{\partial}{\partial q^a},\frac{\partial}{\partial q^b}\right)=R^i_{ab}\frac{\partial}{\partial q^i},$$

where

$$R^i_{ab}=\frac{\partial\Gamma^i_a}{\partial q^b}-\frac{\partial\Gamma^i_b}{\partial q^a}+\Gamma^j_a\frac{\partial\Gamma^i_b}{\partial q^j}-\Gamma^j_b\frac{\partial\Gamma^i_a}{\partial q^j}.$$

We say that Γ is flat if the curvature R identically vanishes. In this case, the constrained system is holonomic.

Notice that this kind of non-holonomic constrained systems is very special, since the local constraints are of the form

$$v^i=-\Gamma^i_a(q^A)v^a,$$

that is, some velocities are explicitly written in terms of the others.

We will consider a very special kind of such constrained systems, those called generalized Čaplygin systems.

Definition IV.1: A generalized Čaplygin system consists of a Lagrangian function $L:TQ\to\mathbb{R}$ and a connection Γ in a fibration $\rho:Q\to M$ such that

$$L((Y^H)_q)=L((Y^H)_{\tilde q}),$$

for any $Y\in T_yM$, where $q,\tilde q\in Q$ are such that $\rho(q)=\rho(\tilde q)=y$.

Remark IV.2: Notice that the Čaplygin systems considered by Koiller[18] are particular cases. In fact, in that case, $\rho:Q\to M$ is a principal bundle with structure group G, L is G-invariant, and Γ is a principal connection, i.e., the horizontal subspaces are G-invariant. The 1-forms η_i are just the components of the connection form. If the group G is Abelian, then the last condition implies that the Christoffel components do not depend on the fiber coordinates. So, we recover the classical setting of Čaplygin systems[18].

From the definition, one easily see that there exists a well-defined Lagrangian function $L^*:TM\to\mathbb{R}$, by setting

$$L^*(Y)=L((Y^H)_q),$$

for any $Y\in T_yM$, where q is an arbitrary point in the fiber over y. In local coordinates we have

$$L^*(q^a,v^a)=L(q^a,q^i,v^a,-\Gamma^i_av^a).$$

Since L^* does not depend on q^i we deduce that

$$\frac{\partial L}{\partial q^i}=\frac{\partial L}{\partial v^j}\frac{\partial\Gamma^j_a}{\partial q^i}v^a. \tag{13}$$

The constrained Euler–Lagrange equations for L are the following:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial v^a}\right)-\frac{\partial L}{\partial q^a}=-\sum_i\lambda^i\Gamma^i_a,$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial v^i}\right)-\frac{\partial L}{\partial q^i}=-\lambda^i,$$

$$v^a=\frac{dq^a}{dt},\quad v^i=\frac{dq^i}{dt}.$$

After some calculations, and using (13), we obtain that

$$\frac{d}{dt}\left(\frac{\partial L^*}{\partial v^a}\right)-\frac{\partial L^*}{\partial q^a}=\frac{\partial L}{\partial v^i}v^b\left(\frac{\partial \Gamma_b^i}{\partial q^a}-\frac{\partial \Gamma_a^i}{\partial q^b}+\Gamma_b^j\frac{\partial \Gamma_a^i}{\partial q^j}-\Gamma_a^j\frac{\partial \Gamma_b^i}{\partial q^j}\right),$$

where $v^a=dq^a/dt$.

As we have proved in Section II, the intrinsic motion equations are

$$\begin{gathered}(i_X\omega_L-dE_L)\in(H^v)^0,\\ X\in H^c,\end{gathered}\tag{14}$$

along H.

If we assume that the constrained system is regular (for instance, if the Lagrangian L is natural) then there exists an almost product structure $(\mathscr{P},\mathscr{Q})$ on TQ along H such that the vector field $\xi=\mathscr{P}(\xi_L)$ gives the constrained dynamics. Let us recall that ξ is a vector field defined on H, that is, $\xi\in\mathfrak{X}(H)$.

Define a 1-form $\alpha_{L,\Gamma}$ on TM as follows:

$$(\alpha_{L,\Gamma})_u(U)=-(\alpha_L)_x(\widetilde{X}),$$

for any $U\in T_u(TM)$, for any $u\in T_yM$, where $\widetilde{X}\in T_x(TQ)$ is a tangent vector which projects onto the tangent vector $R((u^H)_q,(T\tau_M(U))_q^H)\in T_qQ$, $\rho(q)=y$, and $x\in D$ with $\tau_Q(x)=q$. [Notice that there is a unique point $x\in D$ such that $\tau_Q(x)=q$ and $T\rho(x)=u$.] In local coordinates, we get

$$\alpha_{L,\Gamma}=\left[\frac{\partial L}{\partial v^i}v^bR_{ab}^i\right]dq^a.$$

It should be remarked that $\alpha_{L,\Gamma}$ is not a *bona fide* 1-form on TM, but it is a 1-form along the mapping $T\rho_{/H}:H\to TM$. For the sake of simplicity, we will assume that $\alpha_{L,\Gamma}$ is well-defined on TM, which is the case in most of the examples.

Now, consider the non-conservative Lagrangian system with Lagrangian function L and external force $\alpha_{L,\Gamma}$. The intrinsic motion equation is

$$i_Y\omega_{L*}=dE_{L*}+\alpha_{L,\Gamma},\tag{15}$$

on TM. We will study its solutions. Notice that the corresponding Euler–Lagrange equations are

$$\begin{gathered}\frac{d}{dt}\left(\frac{\partial L^*}{\partial v^a}\right)-\frac{\partial L^*}{\partial q^a}=-\frac{\partial L}{\partial v^i}v^bR_{ab}^i,\\ v^a=\frac{dq^a}{dt}.\end{gathered}\tag{16}$$

Theorem IV.3: *(1) The generalized Čaplygin system* (L,Γ) *is regular* iff L^* *is regular; (2) In this case, the vector field* ξ *is projectable onto* TM*, and its projection* Y *is just the solution of (15).*

Dynamical proof: We first show how the result can be derived by using a pure dynamical argument. As we have shown, if $(q^a(t),q^i(t))$ is a solution of the constrained motion equations (14) then it is a horizontal curve, and its projection is a solution of the non-conservative equations (16). Conversely, if $(q^a(t))$ is a solution of (16), then its horizontal lift to Q is a solution of (14). Now, assume that the generalized Čaplygin system (L,Γ) is regular so that there exists one and only one solution with a fixed initial data in TM. Take an initial data in TM. Its horizontal lift gives an initial data in TQ for which there exists one and only one solution of (14). Its projection will be a solution of (15) for the given initial data, and, furthermore, it will be the unique solution with that data. The converse is proved by a similar argument.

The same procedure proves that ξ is projectable, and its projection Y is just a solution of (15). In fact, the solutions of ξ project onto the solutions of Y, and, conversely, the horizontal lifts of the solutions of Y are just the solutions of ξ.

Next, we exhibit an alternative proof based on the geometrical ingredients of the theory. First of all, we will prove the following lemma.

Lemma IV.4: Let Γ *be an arbitrary connection in a fibration* $\rho:Q\rightarrow M$ *with horizontal projector* $\mathbf{h}$*. If* μ_1 *and* μ_2 *are two 1-forms and* X *is a horizontal vector field on* Q *such that*

$$\mathscr{L}_X\mu_1=\mu_2,$$

then we have

$$\mathscr{L}_X(\mathbf{h}^*\mu_1)=\mathbf{h}^*\mu_2-\alpha,$$

where $\mathbf{h}^*$ *is the transpose operator of* $\mathbf{h}$*, and* α *is the 1-form on* Q *defined by*

$$\alpha(Y)=-\mu_1(R(X,Y)-\mathbf{h}([X,\mathbf{v}Y])),$$

R *being the curvature of* Γ.

Proof: Assume that $\mathscr{L}_X\mu_1=\mu_2$ and X is a horizontal vector field. Let Y be an arbitrary vector field on Q. We have

$$\begin{aligned}(\mathscr{L}_X(\mathbf{h}^*\mu_1))(Y)&=\mathscr{L}_X(\mathbf{h}^*\mu_1)(Y)-(\mathbf{h}^*\mu_1)([X,Y])\\&=X(\mu_1(\mathbf{h}Y))-\mu_1(\mathbf{h}([X,Y])\\&=\mu_2(\mathbf{h}Y)+\mu_1([X,\mathbf{h}Y])-\mu_1(\mathbf{h}[X,Y])\\&=\mathbf{h}^*\mu_2(Y)+\mu_1(R(X,Y)-\mathbf{h}([X,\mathbf{v}Y])).\end{aligned}$$

■

Geometrical proof of the theorem: First of all, we will prove that the generalized Čaplygin system (L,Γ) is regular if L^* is regular. In fact, denote by

$$W=(W^{AB})=\begin{pmatrix}W^{ab} & W^{aj}\\ W^{ib} & W^{ij}\end{pmatrix}$$

the inverse matrix of the Hessian matrix $(\partial^2L/\partial v^A\partial v^B)$. We have

$$\mathscr{C}_{ij}=-W^{ab}\Gamma_a^i\Gamma_b^j-W^{ja}\Gamma_a^i-W^{ib}\Gamma_b^j-W^{ij},$$

or, equivalently,

$$\mathscr{C}=-(\gamma,I_{m\times m})W(\gamma,I_{m\times m})^t,$$

where γ is a matrix $m\times(n-m)$ with entries $\gamma_{ia}=\Gamma_a^i$, $1\leqslant i\leqslant m$, $1\leqslant a\leqslant n-m$, and the superindex t means that we are taking the transpose matrix.

On the other hand, the entries of the Hessian matrix $\mathscr{M}$ of L^* are

$$\frac{\partial^2L^*}{\partial v^a\partial v^b}=\frac{\partial^2L}{\partial v^a\partial v^b}-\Gamma_a^i\frac{\partial^2L}{\partial v^i\partial v^b}-\Gamma_b^j\frac{\partial^2L}{\partial v^j\partial v^a}+\Gamma_a^i\Gamma_b^j\frac{\partial^2L}{\partial v^i\partial v^j},$$

or, equivalently,

$$\mathscr{M}=(I_{(n-m)\times(n-m)},-\gamma^t)W^{-1}(I_{(n-m)\times(n-m)},-\gamma^t)^t.$$

If we put $\mathscr{A}=(\gamma,I_{m\times m})$ and $\mathscr{B}=(I_{(n-m)\times(n-m)},-\gamma^t)$, it is no hard to show that either $\mathscr{C}$ or $\mathscr{M}$ are regular, then the square matrices

$$\begin{pmatrix} \mathscr{A} \\ \mathscr{B}W^{-1} \end{pmatrix} \text{ and } \begin{pmatrix} W\mathscr{A}^t \\ \mathscr{B}^{\,t} \end{pmatrix}$$

are also regular.

The result follows taking into account that

$$\begin{pmatrix} \mathscr{A} \\ \mathscr{B}W^{-1} \end{pmatrix}\cdot\begin{pmatrix} W\mathscr{A}^t \\ \mathscr{B}^t \end{pmatrix}=\begin{pmatrix} \mathscr{A}W\mathscr{A}^t & 0 \\ 0 & \mathscr{B}W^{-1}\mathscr{B}^{\,t} \end{pmatrix}.$$

Therefore, we have proved the first part of the theorem.

Next, we will prove the second part.

Given a connection Γ in the fibration $\rho:Q\rightarrow M$ we define a connection $\overline{\Gamma}$ in the fibration $T\rho:TQ\rightarrow TM$ along the submanifold H as follows. The horizontal distribution $\overline{H}$ of $\overline{\Gamma}$ is locally spanned by the vector fields

$$\left(\frac{\partial}{\partial q^a}\right)^{\overline{H}}=\frac{\partial}{\partial q^a}-\Gamma_a^i\frac{\partial}{\partial q^i}-v^b\left(\frac{\partial\Gamma_b^i}{\partial q^a}-\Gamma_a^j\frac{\partial\Gamma_b^i}{\partial q^j}\right)\frac{\partial}{\partial v^i},$$

$$\left(\frac{\partial}{\partial v^a}\right)^{\overline{H}}=\frac{\partial}{\partial v^a}-\Gamma_a^i\frac{\partial}{\partial v^i}.$$

Along H, we obtain a local basis of vector fields on TQ,

$$\left\{\left(\frac{\partial}{\partial q^a}\right)^{\overline{H}},\left(\frac{\partial}{\partial v^a}\right)^{\overline{H}},\frac{\partial}{\partial q^i},\frac{\partial}{\partial v^i}\right\}.$$

Its dual basis of 1-forms is

$$\{dq^a,dv^a,\eta_i^v,d\hat{\eta}_i\}.$$

Thus, the set $\{\eta_i^v,d\hat{\eta}_i\}$ is the annihilator of $\overline{H}$. A simple computation shows that $\overline{H}$ is globally defined along H.

If $\overline{\mathbf{h}}$ is the horizontal projector associated with $\overline{\Gamma}$ we have $\overline{\mathbf{h}}^*(dq^a)=dq^a$, $\overline{\mathbf{h}}^*(dv^a)=dv^a$, $\overline{\mathbf{h}}^*(\eta_i^v)=0$ and $\overline{\mathbf{h}}^*(d\hat{\eta}_i)=0$.

Consider the pull-backs of the 1-forms α_{L*} and $dL*$ to TQ by means of $T\rho$. Along H we deduce that

$$\overline{\mathbf{h}}^*(\alpha_L)=(T\rho)^*\alpha_{L*},$$

$$\overline{\mathbf{h}}^*(dL)=(T\rho)^*dL*.$$

From Lemma II.8 we have

$$\mathscr{L}_\xi\alpha_L=dL-\mathscr{L}_{\mathscr{Q}(\xi_L)}\alpha_L, \tag{17}$$

and from Lemma IV.4, we get

$$\mathscr{L}_{\xi}(\overline{\mathbf{h}^{*}}\alpha_L)=\overline{\mathbf{h}^{*}}(dL)-\overline{\mathbf{h}^{*}}(\mathscr{L}_{\mathscr{Q}(\xi_L)}\alpha_L)-\overline{\alpha}, \tag{18}$$

where $\overline{\alpha}$ is the 1-form on TQ along H defined by

$$\overline{\alpha}(Z)=-\alpha_L(\overline{R}(\xi,Z)-\overline{\mathbf{h}}([\xi,\overline{\mathbf{v}}(Z)])),$$

$\overline{R}$ being the curvature of $\overline{\Gamma}$. Since α_L is semibasic and $\overline{\Gamma}$ is a connection in the fibration $T\rho{:}TQ\rightarrow TM$ (along H), we deduce that $\alpha_L(\overline{\mathbf{h}}([\xi,\overline{\mathbf{v}}(Z)]))=0$, and hence we get

$$\overline{\alpha}(Z)=-\alpha_L(\overline{R}(\xi,Z)).$$

In local coordinates we obtain

$$\overline{\alpha}=\frac{\partial L}{\partial v^i}v^b\left(\frac{\partial\Gamma^i_a}{\partial q^b}-\frac{\partial\Gamma^i_b}{\partial q^a}-\Gamma^j_b\frac{\partial\Gamma^i_a}{\partial q^j}+\Gamma^j_a\frac{\partial\Gamma^i_b}{\partial q^j}\right)dq^a.$$

Therefore, we deduce that $\overline{\alpha}$ is projectable, and its projection is just the 1-form $\alpha_{L,\Gamma}$ on TM.

From Lemma II.9 we have

$$\overline{\mathbf{h}^{*}}(\mathscr{L}_{\mathscr{Q}(\xi_L)}\alpha_L)=0,$$

and therefore (18) becomes

$$\mathscr{L}_{\xi}(T\rho)^{*}\alpha_{L^{*}}=(T\rho)^{*}(dL^{*})-\overline{\alpha}.$$

Let Y be a vector field on TM which is a solution of the equation

$$\mathscr{L}_{Y}\alpha_{L^{*}}=dL^{*}-\alpha_{L,\Gamma}.$$

Then every vector field $\widetilde{Y}$ on TQ which projects onto Y verifies

$$\mathscr{L}_{\widetilde{Y}}(T\rho)^{*}\alpha_{L^{*}}=(T\rho)^{*}(dL^{*})-\overline{\alpha}. \tag{19}$$

In particular, the horizontal lift $Y^{\overline{H}}$ with respect to $\overline{\Gamma}$ verifies (19). Since ξ also verifies (19) and $\xi\in\overline{H}$, we deduce that $Y^{\overline{H}}=\xi$.

■

Thus, we have the following.

Corollary IV.5: The generalized Čaplygin system (L,Γ) is equivalent to a non-conservative system on TM with Lagrangian function L^{*} and external force $\alpha_{L,\Gamma}$.

Remark IV.6: The above procedure is a sort of reduction, but not in the sense of Marsden and Weinstein.[40] In fact, we could consider the general case of a constrained Lagrangian system subjected to linear constraints given by a distribution D on Q, and such that L and D are invariant by the action of a Lie group G. This is just the case of Čaplygin systems as were considered by Koiller.

Remark IV.7: The distribution H^c satisfies the following relation:

$$T(TQ)=H^c\oplus V(T\rho),$$

along the points of H. Thus, H^c defines a connection Γ^c in the fibration $T\rho{:}TQ\rightarrow TM$ along the submanifold H, which could be considered as the tangent lift of the original connection Γ in the fibration $\rho{:}Q\rightarrow M$. We have proved that the vector field ξ is horizontal with respect to $\overline{\Gamma}$. A similar device proves that ξ is also horizontal with respect to Γ^c, and moreover $\xi = Y^{H^c}$. The relationship between both connections $\overline{\Gamma}$ and Γ^c is the following: they coincide if and only Γ is flat, or, in other words, the constrained system is holonomic.

Remark IV.8: Assume that the constrained system is not regular. From Theorem IV.3 we deduce that L^* is a singular Lagrangian function. Thus, Equation (15) has no, in general, solution. However, we can develop a constraint algorithm as follows. Put $K_1 = TM$ and define K_2 be the submanifold of points in TM for which there exists at least a solution of (15). On K_2 there is a solution, but it is not necessarily tangent to K_2. So, we consider the submanifold K_3 consisting in those points in K_2 where a tangent solution to K_2 exists.

Proceeding further we obtain a sequence of constraint submanifolds,

$$\cdots \to K_k \to \cdots \to K_2 \to K_1 = TM.$$

On the other hand, there exists a sequence of constraint submanifolds,

$$\cdots \to H_k \to \cdots \to H_2 \to H_1 = H,$$

obtained by applying the constraint algorithm developed in Section III. It is almost obvious that both algorithms are related by the projection mapping $T\rho: TQ \to TM$, that is, we have

$$T\rho(H_r) = M_r, \quad r \geqslant 1.$$

Example IV.9 (The sleigh of Čaplygin and Carathéodory):

Consider a sleigh, that is, a body having three points of contact with a plane where two of them slide freely but the third A is subjected to a force which does not allow transversal velocity. The configuration manifold is $Q = \mathbb{R}^2 \times S^1$ with coordinates (x,y,ϕ), where (x,y) are the coordinates of the center of mass C of the sleigh, and ϕ is the angle between the x-axis and the line AC (see Ref. 18). If we denote by a the distance from A to C, by J the moment of inertia and by $m=1$ the mass of the sleigh, the Lagrangian function is given by

$$L = \frac{1}{2}(\dot{x}^2 + \dot{y}^2) + \frac{1}{2} J \dot{\phi}^2,$$

Observe that L is a natural Lagrangian obtained from the Riemannian metric

$$g = \frac{1}{2}(dx^2 + dy^2 + J d\phi^2),$$

on $\mathbb{R}^2 \times S^1$.

Consider the fibration

$$\rho: \mathbb{R}^2 \times S^1 \to \mathbb{R}^2,$$

defined by

$$\rho(x,y,\phi) = (x,y).$$

Define a connection Γ in ρ by

$$\Gamma\left(\frac{\partial}{\partial x}\right)=\frac{\partial}{\partial x}-2\frac{\sin\phi}{a}\frac{\partial}{\partial\phi},$$

$$\Gamma\left(\frac{\partial}{\partial y}\right)=\frac{\partial}{\partial y}+2\frac{\cos\phi}{a}\frac{\partial}{\partial\phi},$$

$$\Gamma\left(\frac{\partial}{\partial\phi}\right)=-\frac{\partial}{\partial\phi}.$$

The curvature R of Γ is

$$R=\frac{1}{a^2}\frac{\partial}{\partial\phi}\otimes(dx\wedge dy).$$

The horizontal distribution of Γ is generated by

$$\left\langle\frac{\partial}{\partial x}-\frac{\sin\phi}{a}\frac{\partial}{\partial\phi},\quad\frac{\partial}{\partial y}+\frac{\cos\phi}{a}\frac{\partial}{\partial\phi}\right\rangle,$$

and the annihilator of H is generated by the 1-form

$$\eta=d\phi-\frac{\cos\phi}{a}dy+\frac{\sin\phi}{a}dx.$$

In fact, η is the connection 1-form of Γ. Therefore, the linear constraints are

$$\dot{\phi}-\frac{\cos\phi}{a}\dot{y}+\frac{\sin\phi}{a}\dot{x}=0.$$

Notice that ρ is a principal S^1-bundle. However, Γ is no a principal connection, since the horizontal subspaces are not S^1-invariant. Thus, (L,Γ) is not a generalized Čaplygin system. However, we can apply the general procedure developed in Section II.

Since L is natural, the constrained system is regular, and then there exists a well-defined solution of the constrained dynamics along the submanifold H of TQ.

The distribution S is generated by the vector field

$$Z=-\frac{1}{J}\frac{\partial}{\partial\dot{\phi}}+\frac{\cos\phi}{a}\frac{\partial}{\partial\dot{y}}-\frac{\sin\phi}{a}\frac{\partial}{\partial\dot{x}},$$

along the points of H.

The almost product structure $(\mathscr{P},\mathscr{Q})$ is given by

$$\mathscr{Q}=-\frac{Ja^2}{a^2+J}\left(-\frac{1}{J}\frac{\partial}{\partial\dot{\phi}}+\frac{\cos\phi}{a}\frac{\partial}{\partial\dot{y}}-\frac{\sin\phi}{a}\frac{\partial}{\partial\dot{x}}\right)\otimes\left(\left(\frac{\sin\phi}{a}\dot{y}+\frac{\cos\phi}{a}\dot{x}\right)d\phi+\frac{\sin\phi}{a}d\dot{x}-\frac{\cos\phi}{a}d\dot{y}+d\dot{\phi}\right),$$

$$\mathscr{P}=\mathrm{id}-\mathscr{Q}.$$

Thus, the only vector field ξ such that $\xi_x \in (S_L)_x \cap T_x H$ and $J_x(\xi_x) = C_x$, for any $x \in H$, is given by

$$\xi = \mathscr{P}(\xi_L) = \dot{x}\frac{\partial}{\partial x} + \dot{y}\frac{\partial}{\partial y} + \dot{\phi}\frac{\partial}{\partial \phi} - \frac{J}{a^2+J}\sin\phi(\sin\phi\dot{y} + \cos\phi\dot{x})\frac{\partial}{\partial \dot{x}}$$

$$+ \frac{J}{a^2+J}\cos\phi(\sin\phi\dot{y} + \cos\phi\dot{x})\frac{\partial}{\partial \dot{y}} - \frac{a}{a^2+J}(\sin\phi\dot{y} + \cos\phi\dot{x})\frac{\partial}{\partial \dot{\phi}},$$

along the points of H.

Example IV.10 (The ''two-wheeled carriage''):

The configuration space of the ''two-wheeled carriage'' is $Q = \mathbb{R}^2 \times S^1 \times T^2$, with coordinates $(x, y, \phi, \Phi_1, \Phi_2)$ (see Refs. 18,3).

Let $2r$ be the lateral length, a the radius of the wheels, C_0 the center of mass, situated at distance l from a point (x,y). If we denote by m_0 the mass of the body without wheels, k_0 the radius of gyration about the vertical through (x,y), m_1 the mass of a wheel, C the axial moments of inertia and A its moment of inertia about a diameter, then the Lagrangian function is given by

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + m_0 l\dot{\phi}(\dot{y}\cos\phi - \dot{x}\sin\phi) + \frac{1}{2}J\dot{\phi}^2 + \frac{1}{2}C(\dot{\Phi}_1^2 + \dot{\Phi}_2^2),$$

where $m = m_0 + 2m_1$ and $J = m_0 k_0^2 + 2m_1 r^2 + 2A$.

Consider now the fibration

$$\rho: \mathbb{R}^2 \times S^1 \times T^2 \to T^2,$$

defined by $\rho(x, y, \phi, \Phi_1, \Phi_2) = (\Phi_1, \Phi_2)$.

Define a connection Γ in ρ by

$$\Gamma\left(\frac{\partial}{\partial \Phi_1}\right) = \frac{\partial}{\partial \Phi_1} - a\cos\phi\frac{\partial}{\partial x} - a\sin\phi\frac{\partial}{\partial y} - \frac{a}{r}\frac{\partial}{\partial \phi},$$

$$\Gamma\left(\frac{\partial}{\partial \Phi_2}\right) = \frac{\partial}{\partial \Phi_2} - a\cos\phi\frac{\partial}{\partial x} - a\sin\phi\frac{\partial}{\partial y} + \frac{a}{r}\frac{\partial}{\partial \phi},$$

$$\Gamma\left(\frac{\partial}{\partial x}\right) = -\frac{\partial}{\partial x}, \quad \Gamma\left(\frac{\partial}{\partial y}\right) = -\frac{\partial}{\partial y}, \quad \Gamma\left(\frac{\partial}{\partial \phi}\right) = -\frac{\partial}{\partial \phi}.$$

The horizontal distribution H of Γ is generated by

$$\left\langle \frac{\partial}{\partial \Phi_1} - \frac{a}{2}\cos\phi\frac{\partial}{\partial x} - \frac{a}{2}\sin\phi\frac{\partial}{\partial y} - \frac{a}{2r}\frac{\partial}{\partial \phi}, \quad \frac{\partial}{\partial \Phi_2} - \frac{a}{2}\cos\phi\frac{\partial}{\partial x} - \frac{a}{2}\sin\phi\frac{\partial}{\partial y} + \frac{a}{2r}\frac{\partial}{\partial \phi} \right\rangle.$$

Thus, the annihilator of H is generated by the 1-forms

$$\eta_x = dx + \frac{a\cos\phi}{2}d\Phi_1 + \frac{a\cos\phi}{2}d\Phi_2,$$

$$\eta_y = dy + \frac{a\sin\phi}{2}d\Phi_1 + \frac{a\sin\phi}{2}d\Phi_2,$$

$$\eta_\phi = d\phi + \frac{a}{2r} d\Phi_1 - \frac{a}{2r} d\Phi_2 .$$

The linear constraints are then

$$\dot{x} = -\frac{a\cos\phi}{2}\dot{\Phi}_1 - \frac{a\cos\phi}{2}\dot{\Phi}_2 ,$$

$$\dot{y} = -\frac{a\sin\phi}{2}\dot{\Phi}_1 - \frac{a\sin\phi}{2}\dot{\Phi}_2 ,$$

$$\dot{\phi} = -\frac{a}{2r}\dot{\Phi}_1 + \frac{a}{2r}\dot{\Phi}_2 .$$

The curvature of the connection is given by the vector 2-form,

$$R = \frac{a^2}{2r}\left(-\sin\phi \frac{\partial}{\partial x} + \cos\phi \frac{\partial}{\partial y} \right) \otimes (d\Phi_1 \wedge d\Phi_2).$$

The fibration $\rho:\mathbb{R}^2\times S^1\times T^2 \to T^2$ is a principal $\mathbb{R}^2\times S^1$-bundle, and Γ is a principal connection with connection 1-form

$$\eta = (\eta_x , \eta_y , \eta_\phi),$$

taking values into the Lie algebra of $\mathbb{R}^2\times S^1$. Notice that $\mathbb{R}^2\times S^1$ may be identified with the group of Euclidean motions of the plane (see Koiller[18]).

Thus, the system (L,Γ) is a generalized Caplygin system. By applying the general theory developed in this section, we obtain a Lagrangian function $L^*:T^2 \to \mathbb{R}$ as follows:

$$L^*(\Phi_1,\Phi_2,\dot{\Phi}_1,\dot{\Phi}_2) = \frac{1}{8} ma^2(\dot{\Phi}_1^2 + \dot{\Phi}_2^2) + \frac{Ja^2}{8r^2}(\dot{\Phi}_2 - \dot{\Phi}_1)^2 + \frac{1}{2} C(\dot{\Phi}_1^2 + \dot{\Phi}_2^2).$$

From Corollary IV.5, we know that the constrained system (L,D) is equivalent to the non-conservative system given by L^* and the external force

$$\alpha_{L,\Gamma} = \frac{m_0 l a^3}{4r^2}(\dot{\Phi}_2 - \dot{\Phi}_1)\dot{\Phi}_2 d\Phi_1 - \frac{m_0 l a^3}{4r^2}(\dot{\Phi}_2 - \dot{\Phi}_1)\dot{\Phi}_1 d\Phi_2 .$$

From a tedious but straightforward computation we have that the solution Y of the equation

$$i_Y \omega_{L^*} = dE_{L^*} + \alpha_{L,\Gamma} ,$$

is the vector field,

$$Y = \dot{\Phi}_1 \frac{\partial}{\partial \Phi_1} + \dot{\Phi}_2 \frac{\partial}{\partial \Phi_2} + K_1(\dot{\Phi}_1 - \dot{\Phi}_2)(K_2\dot{\Phi}_2 - Ja^2\dot{\Phi}_1)\frac{\partial}{\partial \dot{\Phi}_1} + K_1(\dot{\Phi}_2 - \dot{\Phi}_1)(K_2\dot{\Phi}_1 - Ja^2\dot{\Phi}_2)\frac{\partial}{\partial \dot{\Phi}_2},$$

where $K_1 = m_0 l a^3/(m^2a^4r^4 + 8ma^2r^4C + 2ma^4r^2J + 16r^4C^2 + 8JCa^2r^2)$ and $K_2 = ma^2r^2 + 4Cr^2 + Ja^2$.

We obtain the solution on TQ by taking the horizontal lift of the vector field Y by the connection $\overline{\Gamma}$.

V. AN APPLICATION: EQUATIONS OF CONSTRAINED GEODESICS

Let Q be a Riemannian manifold with Riemannian metric g and Levi-Civita connection ∇ and suppose that a distribution D on Q is given. A very old problem in the literature is to obtain a new linear connection ∇^* on Q such that the geodesics of ∇^* are the extremals of the variational problem subjected to these linear constraints (see Synge,[35] Vranceanu,[36] Neimark and Fufaev[3] and the references therein.) We shall apply our method to give a new look at Synge's paper.

The Lagrangian function is

$$L(q^A,v^A)=\frac{1}{2}g_{AB}v^Av^B,$$

that is, L is the kinetic energy of g. Take an orthonormal local basis $\{\mu_i\}$ of D. Since $E_L=L$, we obtain

$$\xi_L=v^A\frac{\partial}{\partial q^A}-\Gamma^C_{AB}v^Av^B\frac{\partial}{\partial v^C},$$

where Γ^C_{AB} are the Christoffel components of ∇. In fact, ξ_L is the geodesic spray.

A direct computation shows that

$$Z_i=-g^{AB}(\mu_i)_B\frac{\partial}{\partial v^A},$$

$$\hat{\mu}_i=(\mu_i)_Av^A,$$

$$\mathscr{C}_{ij}=Z_i(\hat{\mu}_j)=-g^{AB}(\mu_i)_A(\mu_j)_B=-\delta_{ij}.$$

Therefore, the constrained system is regular.

From Proposition II.4 there exists a unique almost product structure $(\mathscr{P},\mathscr{Q})$ such that $\mathscr{P}_x(X)\in T_xD$ and $\mathscr{Q}_x(X)\in S_x$, where $X\in T_xTQ$. We have

$$\begin{aligned}\mathscr{P}(\xi_L)&=v^A\frac{\partial}{\partial q^A}-v^Av^B\left(\Gamma^C_{AB}+\frac{\partial(\mu_i)_B}{\partial q^A}g^{CR}(\mu_i)_R-\Gamma^E_{AB}(\mu_i)_Eg^{CR}(\mu_i)_R\right)\frac{\partial}{\partial v^C}\\&=v^A\frac{\partial}{\partial q^A}-v^Av^B\left(\Gamma^C_{AB}+\left(\frac{\partial(\mu_i)_B}{\partial q^A}-\Gamma^E_{AB}(\mu_i)_E\right)g^{CR}(\mu_i)_R\right)\frac{\partial}{\partial v^C}\\&=v^A\frac{\partial}{\partial q^A}-v^Av^B(\Gamma^C_{AB}+(\mu_i)_{A;B}(\mu_i)^C)\frac{\partial}{\partial v^C},\end{aligned}$$

where

$$(\mu_i)_{A;B}=\frac{\partial(\mu_i)_A}{\partial q^B}-\Gamma^E_{AB}(\mu_i)_E$$

denote the components of the covariant derivative of μ_i, and $(\mu_i)^C=g^{CR}(\mu_i)_R$.

Since $\mathscr{P}(\xi_L)$ is a SODE and tangent to D, we know that for each tangent vector $z\in D$ there is a curve σ on Q which is a solution of $\mathscr{P}(\xi_L)$ with that initial data, i.e., $\sigma(0)=x$, $\dot{\sigma}(0)=z$ and $\dot{\sigma}$ is an integral curve of $\mathscr{P}(\xi_L)$. In fact, the solutions of $\mathscr{P}(\xi_L)$ are just the solutions of the following system of differential equations:

$$\frac{d^2q^C}{dt^2}+(\Gamma^*)^C_{AB}\frac{dq^A}{dt}\frac{dq^B}{dt}=0, \tag{20}$$

where

$$(\Gamma^*)^C_{AB}=\Gamma^C_{AB}+(\mu_i)_{A;B}(\mu_i)^C.$$

Notice that (20) are just the differential equations obtained by Synge in Ref. 35. Of course, (20) have no solutions for arbitrary initial data, since only the velocities in D are allowable.

Notice that $(\Gamma^*)^A_{BC}$ are not the Christoffel components of a linear connection ∇^* on Q. They define a more general geometric object, an spray defined on a submanifold D of TQ. Indeed, let us recall that there exists a one-to-one correspondence between sprays on TQ and linear connections on Q (see Ref. 38). In fact, if ξ is an spray, then $\Gamma=-\mathscr{L}_\xi J$ is a linear connection on Q, and, conversely, if Γ is a linear connection on Q, then its associated SODE is an spray.

The vector field $\mathscr{P}(\xi_L)$ can be extended to a vector field defined on some open neighborhood of D in TQ. Of course, there are many extensions of $\mathscr{P}(\xi_L)$. Choose an arbitrary extension and define

$$\Gamma^*=-\mathscr{L}_{P(\xi_L)}J.$$

So, Γ^* is a tensor field of type (1,1) defined on some open neighborhood of D, and its local expression is as follows:

$$\Gamma^*\left(\frac{\partial}{\partial q^A}\right)=\frac{\partial}{\partial q^A}-2(\Gamma^*)^C_{AB}v^B\frac{\partial}{\partial v^C},$$

$$\Gamma^*\left(\frac{\partial}{\partial v^A}\right)=-\frac{\partial}{\partial v^A}.$$

A direct computation shows that $(\Gamma^*)^2=\mathrm{id}$, and the vector eigenspace corresponding to the eigenvalue -1 at a point of D is just the vertical subspace at that point. Moreover, given another extension of $\mathscr{P}(\xi_L)$, we obtain that the new tensor field Γ^* coincides with the former on D.

Thus, Γ^* defines a connection on some open neighborhood of D and all these connections coincide on D.

We define the horizontal and vertical projectors of Γ^* in the usual way:

$$h^*=\frac{1}{2}(\mathrm{id}+\Gamma^*),\quad v^*=\frac{1}{2}(\mathrm{id}-\Gamma^*).$$

Their local expressions are the following ones:

$$h^*\left(\frac{\partial}{\partial q^A}\right)=\frac{\partial}{\partial q^A}-(\Gamma^*)^C_{BA}v^B\frac{\partial}{\partial v^C},$$

$$h^*\left(\frac{\partial}{\partial v^A}\right)=0,$$

$$v^*\left(\frac{\partial}{\partial q^A}\right)=(\Gamma^*)^C_{BA}v^B\frac{\partial}{\partial v^C},$$

$$v^*\left(\frac{\partial}{\partial v^A}\right)=\frac{\partial}{\partial v^A}.$$

Using the standard procedures for connections on TQ (see Refs. 37,38), we define a covariant derivative as follows. Let X be a vector field on Q and Y be a vector field which is tangent to D. In other words, X is a section of $\tau_Q : TQ \to Q$ and Y a section of $\tau_{Q/D} : D \to Q$. Define

$$(\nabla^*_X Y)(x) = \phi_{Y(x)}(v^*(dY(x)(X(x)))), \quad \forall x \in Q,$$

where $\phi_{Y(x)}$ is the linear isomorphism,

$$\phi_{Y(x)} : V_{Y(x)} \tau_Q \to T_x Q$$

from the vertical subspace at $Y(x)$ onto $T_x Q$.

If $X = X^A(\partial/\partial q^A)$ and $Y = Y^A(\partial/\partial q^A)$, we deduce that

$$\nabla^*_X Y = X^A \left[\frac{\partial Y^C}{\partial q^A} + (\Gamma^*)^C_{AB} Y^B \right] \frac{\partial}{\partial q^C}.$$

We look for a condition which ensures that $\nabla^*_X Y \in D$. We have

$$\mu_i(\nabla^*_X Y) = (\mu_i)_D dq^D \left[\left(X^A \frac{\partial Y^C}{\partial q^A} + X^A (\Gamma^*)^C_{AB} Y^B \right) \right]$$
$$= X^A \left[(\mu_i)_C \frac{\partial Y^C}{\partial q^A} + (\mu_i)_C (\Gamma^*)^C_{AB} Y^B \right].$$

Since

$$(\Gamma^*)^A_{BC} = \Gamma^C_{AB} + (\mu_i)_{A;B} (\mu_i)^C,$$

we deduce that

$$(\mu_i)_C (\Gamma^*)^C_{AB} = (\mu_i)_C \Gamma^C_{AB} + (\mu_i)_C \frac{\partial(\mu_J)_A}{\partial q^B} (\mu_j)^C - (\mu_i)^C \Gamma^E_{AB} (\mu_i)_E (\mu_j)^C$$
$$= \frac{\partial(\mu_i)_A}{\partial q^B}.$$

Thus, we obtain that

$$\mu_i(\nabla^*_X Y) = X^A \left[(\mu_i)_B \frac{\partial Y^B}{\partial q^A} + \frac{\partial(\mu_i)_A}{\partial q^B} Y^B \right]. \tag{21}$$

But $Y \in D$, and therefore we get

$$0 = \mu_i(Y) = (\mu_i)_B Y^B.$$

By deriving this last formula, we have

$$(\mu_i)_B \frac{\partial Y^B}{\partial q^A} + \frac{\partial(\mu_i)_B}{\partial q^A} Y^B = 0. \tag{22}$$

From (21) and (22) we deduce the following result.

Proposition V.1: ∇^* defines a connection in the vector bundle $\tau_Q : D \to Q$ if and only if D is involutive.

As a consequence, if the system is holonomic, ∇^* is a derivation in the vector bundle $D \to Q$. In the general case, we only get that

$$\nabla^*: \mathfrak{X}(Q) \times \mathrm{Sec}(D) \to \mathfrak{X}(Q)$$

behaves as a derivation.

ACKNOWLEDGMENTS

This work has been partially supported through Grant No. DGICYT (Spain), Project No. PB94-0106. The authors acknowledge F. Cantrijn and W. Sarlet for several useful conversations.

[1] V. Valcovici, ''Une extension des liaisions non holonomes et des principes variationnels,'' Ber. Verh. Sachs. Akad. Wiss. Leipzig, Math. Nat. Kl. **102**, 1–39 (1958).
[2] L. Pars, *A Treatise on Analytical Dynamics* (Heinemann, New York, 1965).
[3] J. Neimark and N. Fufaev, *Dynamics of Nonholonomic Systems, Transactions of Mathematical Monographs* (AMS, Providence, 1972), Vol. 33.
[4] A. M. Vershik and L. D. Faddeev, ''Differential Geometry and Lagrangian Mechanics with constraints,'' Sov. Phys. Dokl. **17**, 34–36 (1972).
[5] E. J. Saletan and A. H. Cromer, ''A variational principle for non-holonomic systems,'' Am. J. Phys. **38**, 892–897 (1970).
[6] V. V. Rumiantsev, ''On Hamilton's principle for nonholonomic systems,'' PMM **42**, 387-399 (1978); translated as J. Appl. Math. Mech. **42**, 407–419 (1979).
[7] Y. Pironneau, ''Sur les liaisions non linéaires déplacement virtuels à travail nul, conditions de Chetaev,'' *Modern Developments in Analytical Mechanics* Torino, 1982, Vol. II, pp. 671–686.
[8] A. M. Vershik and V. Ya. Gershkovich, ''Nonholonomic dynamical systems, geometry of distributions and variational problems,'' in *Encyclopaedia of Mathematical Sciences*, Dynamical Systems, VII, edited by V. I. Arnold and S. P. Novikov (Springer-Verlag, Berlin, 1994), Vol. 16, pp. 1–81.
[9] E. Massa and E. Pagani, ''Classical dynamics of non-holonomic systems: A geometric approach,'' Ann. Inst. Henri Poincaré: Phys. Theor. **55**, 511–544 (1991).
[10] E. Massa and E. Pagani, ''A new look at Classical Mechanics of constrained systems,'' preprint, 1995.
[11] R. M. Rosenberg, *Analytical Dynamics of Discrete Systems* (Plenum, New York, 1977).
[12] R. Weber, ''Hamiltonian systems with constraints and their meaning in mechanics,'' Arch. Rat. Mech. Anal. **91**, 309–335 (1986).
[13] P. Pitanga and K. C. Mundin, ''Projector in constrained quantum dynamics,'' Nuovo Cimento A **101 A**, 345–352 (1989).
[14] P. Pitanga, ''Symplectic projector in constrained systems,'' Nuovo Cimento A **103 A**, 1529–1533 (1990).
[15] C. M. Marle, ''Sur la géometrie des systèmes mécaniques à liaisions actives,'' C. R. Acad. Sci. Paris **311**, 839–845 (1990).
[16] L. Bates and J. Śniatycki, ''Nonholonomic reduction,'' Rep. Math. Phys. **32**, 99–115 (1992).
[17] G. Giachetta, ''Jet methods in nonholonomic mechanics,'' J. Math. Phys. **33**, 1652–1665 (1992).
[18] J. Koiller, ''Reduction of some classical non-holonomic systems with symmetry,'' Arch. Rat. Mech. Anal. **118**, 113–148 (1992).
[19] J. F. Cariñena and M. F. Rañada, ''Lagrangians systems with constraints: A geometric approach to the method of Lagrange multipliers,'' J. Phys. A **26**, 1335–1351 (1993).
[20] M. F. Rañada, ''Time-dependent Lagrangians systems: A geometric approach to the theory of systems with constraints,'' J. Math. Phys. **35**, 748–758 (1994).
[21] P. Dazord, ''Mécanique Hamiltonienne en présence de constraintes,'' Illinois J. Math. **38**, 148–175 (1994).
[22] J. F. Cariñena and M. F. Rañada, ''Comments on the presymplectic formalism and the theory of regular Lagrangians with constraints,'' J. Phys. A **28** L91–L97 (1995).
[23] W. Sarlet, F. Cantrijn, and D. J. Saunders, ''A geometrical framework for the study of non-holonomic Lagrangian systems,'' J. Phys. A **28**, 3253–3268 (1995).
[24] W. Sarlet, F. Cantrijn, and D. J. Saunders, ''A geometrical framework for the study of non-holonomic Lagrangian systems: II,'' preprint, 1996.
[25] W. Sarlet, ''A direct geometrical construction of the dynamics of non-holonomic Lagrangian systems,'' *Proceedings of the 3rd Meeting on Current Ideas in Mechanics and Related Fields*, Segovia, Spain, 19–23 June 1995, Ext. Math. **10** (1996).
[26] W. Sarlet, ''The geometry of mixed first and second-order differential equations with applications to non-holonomic mechanics,'' in *Proceedings of the 6th International Conference on Differential Geometry and Its Applications*, Brno, August 28–September 1, 1995.
[27] M. de León and D. Martín de Diego, ''Solving non-holonomic Lagrangian dynamics in terms of almost product structures,'' to appear in Ext. Math.
[28] M. de León and D. Martín de Diego, ''A constraint algorithm for singular Lagrangians subjected to non-holonomic constraints,'' preprint, 1995.

[29] M. de León and D. Martín de Diego, ''Non-holonomic mechanical systems in jet bundles,'' in Ref. 25.
[30] M. de León and D. Martín de Diego, ''Almost product structures in Mechanics,'' in Ref. 26.
[31] M. de León and D. Martín de Diego, ''A symplectic formulation of non-holonomic Lagrangian systems,'' in *Proceedings of the IV Fall Workshop: Differential Geometry and its Applications*, Santiago de Compostela, 17–20 September 1995, Anales de Física, Monografías, 1996.
[32] P. A. M. Dirac, *Lecture on Quantum Mechanics*, Belfer Graduate School of Science (Yeshiva University, New York, 1964).
[33] M. J. Gotay and J. M. Nester, ''Presymplectic Lagrangian systems I: The constraint algorithm and the equivalence theorem,'' Ann. Inst. H. Poincaré A **30**, 129–142 (1979).
[34] M. J. Gotay and J. M. Nester, ''Presymplectic Lagrangian systems II: The second-order differential equation problem,'' *Ann. Inst. H. Poincaré* A **32**, 1–13 (1980).
[35] J. L. Synge, ''Geodesics in non-holonomic geometry,'' Math. Ann. **99**, 738–751 (1928).
[36] G. Vranceanu, *Opera Matematica* (3v), Editura Academiei Republicii Socialiste România, Bucharest, 1969.
[37] J. Grifone, ''Structure presque-tangente et connexions, I, II,'' Ann. Inst. Fourier, Grenoble **22**, 287–334 (1972); **22**, 291–338 (1972).
[38] M. de León and P. R. Rodrigues, *Methods of Differential Geometry in Analytical Mechanics* (North-Holland, Amsterdam, 1989).
[39] K. Yano and S. Ishihara, *Tangent and Cotangent Bundles* (Marcel Dekker, New York, 1973).
[40] R. Abraham and J. E. Marsden, *Foundations of Mechanics*, 2nd ed. (Benjamin-Cummings, Reading, MA, 1978).